Common Methodological Challenges Encountered With Multiple Systems Estimation Studies

Article

Common Methodological Challenges Encountered With Multiple Systems Estimation Studies

Crime & Delinquency
1–13

DOI: 10.1177/0011128720981900
journals.sagepub.com/home/cad
SAGE

Kyle Shane Vincent[1], Serveh Sharifi Far[2], and Michail Papathomas[3]

Abstract

Multiple systems estimation refers to a class of inference procedures that are commonly used to estimate the size of hidden populations based on administrative lists. In this paper we discuss some of the common challenges encountered in such studies. In particular, we summarize theoretical issues relating to the existence of maximum likelihood estimators, model identifiability, and parameter redundancy when there is sparse overlap among the lists. We also discuss techniques for matching records when there are no unique identifiers, exploiting covariate information to improve estimation, and addressing missing data. We offer suggestions for remedial actions when these issues/challenges manifest. The corresponding R coding packages that can assist with the analyses of multiple systems estimation data sets are also discussed.

Keywords

covariate information, local MSE challenges, matching records, missing observations, model identifiability

[1]Independent Researcher and Consultant, Ottawa, ON, Canada
[2]The University of Edinburgh School of Mathematics, UK
[3]University of St. Andrews, UK

Corresponding Author:
Kyle Shane Vincent, Independent Researcher and Consultant, Ottawa, ON K2V 0L9, Canada.
Email: kyle.shane.vincent@gmail.com

Introduction

Multiple systems estimation (MSE) is a rapidly growing class of quantification methods that are used for studying hidden populations, such as those comprised of human trafficking victims. The motivation behind MSE is the United Nations (UN) recommendation to monitor the number of victims of human trafficking (per 100,000 population). Victims can be detected or undetected, with data typically collected over a wide span of time. Data sets arise in the form of merged administrative lists, with each list created by a different organization such as the Region or Border police, hospitals, support and protection programs, and non-governmental organizations. As a standard, sex and age, and possibly form of exploitation, are recorded. However, other information can also be collected. For example, in the Netherlands, a data set was collected from a number of sources over 6 years, comprising six lists and five covariates (age, gender, form of exploitation, nationality, and year); see Cruyff et al. (2017) for further information, and see Bird and King (2018) for specific details on how administrative list data are collected.

MSE can be considered a generalization of mark-recapture procedures in that sophisticated mark-recapture modeling of overlaps of "captures" between the administrative lists can be used to estimate the size of the population. Applications of MSE procedures are typically based on a set of capture histories that correspond to the administrative lists. Essentially, lists are first ordered and each individual has a capture history that corresponds to a vector of zeros and ones where these, respectively and keeping with mark-recapture terminology, refer to a "miss" and a "capture" on the corresponding lists. The set of capture histories are concatenated to form a capture history matrix. Much like with mark-recapture procedures (Williams et al., 2002), the capture history matrix forms the data set for which MSE procedures are applied.

When the lists are combined with categorical covariates a number of possible cross-classifications are generated. For example, for the Netherlands data, one cross-classification for an individual observed in the first two lists only could be {Yes, Yes, No, No, No, No} with covariate class {adult, female, beggary, Romania, 2010}. Evidently, the more lists and covariates that are considered the larger the number of possible cross-classifications. Consequently, the larger the probability that no individual is observed for a number of them due to the limited sample size. For instance, in Cruyff et al. (in press), a data set from Slovakia is analyzed where three lists and three covariates (sex, age and type of exploitation) create 64 possible cross-classifications with observed victims in only 21 of them. Such sparseness gives rise to challenges in estimating the number of victims of human trafficking, as discussed in section "Non-Overlapping Lists."

MSE is still a relatively new topic. Original contributions have been made by Bales et al. (2015), where the lists considered are those which arise from the United Kingdom National Crime Agency, and by Cruyff et al. (2017), where lists are based on reports by various organizations to a government-funded NGO, Coordination of Human Trafficking, in the Netherlands. Bird and King (2018) provide a comprehensive summary of MSE methods and applications. Difficulties with model-fitting for MSE data sets have been detailed in Silverman (2020), and novel approaches to model-fitting have been presented to help resolve such issues.

There are several challenges that are commonly encountered in MSE, primarily due to the fact that the study population is comprised of human beings. For example, as people are typically conscious about "self-selection" and/or reporting their identities to multiple administrative lists that are based on a hidden population, challenges arise in modeling the erratic patterns of capture histories. The current mark-recapture literature, which is primarily focused on studying wildlife populations, does not place a focus on such challenges. In this paper we discuss these challenges and summarize methods that can address or account for the limitations in commonly used MSE/mark-recapture estimation procedures.

Non-Overlapping Lists

In the context of MSE, it is not uncommon to observe little to no overlap between administrative lists. This may be due to (1) the fact there is a negative correlation ("trap-shy") effect between pairs of lists; for example, if two lists correspond to service providers that offer similar services to human trafficking victims, then there may not be a need/tendency for individuals to obtain services from more than one, (2) a structural zero; for example, one list may be a service provider that only provides service to females, while another only to males, and/or (3) by chance, which is likely to happen when there are small sample sizes. An example of such a data set, based on data collated from a number of sources in the New Orleans area, is presented in Table 1. In total, 185 individuals are listed as being captured at least once across all administrative sources. Very few individuals are captured more than once, giving rise to a "sparse MSE data set" commonly seen when based on such sources. For further information on this data set and results based on an MSE analysis, see Bales et al. (2018).

Sparsity may lead to difficulty in fitting MSE/mark-recapture models and numerical instability in the resulting estimates. One possible approach is to either combine such pairs of lists into one or to remove the smaller lists altogether (Sharifi Far et al., 2020). However, in doing so there is typically a

Table 1. A Modern Slavery and Trafficking Data Set Based on Several Administrative Sources in New Orleans (Bales et al., 2018).

Cases observed only on one list		Cases observed on exactly two lists		Cases observed on exactly three lists	
List	Number	Lists	Number	Lists	Number
A	25	A&C	1	A&C&G	1
B	5	A&D	2	A&D&E	1
C	70	A&E	1		
D	33	B&F	1		
E	6	C&D	1		
F	6	C&E	1		
G	6	C&G	1		
H	21	D&E	2		
		E&H	1		

Note. List combinations for which no cases are observed are omitted.

reduction or loss of information that can be exploited for inferential purposes. This section considers two approaches that can handle such cases.

Addressing Non-Existence of Maximum Likelihood Estimators and Model Unidentifiability

For MSE applications, one commonly used approach is to fit a Poisson log-linear model to counts of the individuals that are observed on each possible combination of the lists. The Poisson distribution models the number of events occurring in an interval of time or space given a known constant mean rate. For example, consider a data set with two lists referred to as X and Y. Individuals could be observed in only one of the two lists, neither of them, or both of them, which results in four different cross-classifications/combinations. Each variable representing a list has two levels, namely 1 and 0, that respectively indicate whether individuals are identified or not by that list. The number of individuals in each case, n_k, arises independently from a Poisson distribution with a mean of μ_k:

$$n_k \mid \mu_k \sim Poisson(\mu_k), \quad k \in \{00,01,10,11\}.$$

Typically, one models the mean number of individuals in each combination so that

$$\log(\mu_k) = \theta + \theta_i^X + \theta_j^Y + \theta_{ij}^{XY}, \quad k \in \{00,01,10,11\} \quad i,j = 0,1.$$

This is a generalized linear model which in the literature is known as a Poisson log-linear model. In this model, θ is an intercept term associated with the mean count of individuals not observed in any lists, θ_i^X and θ_j^Y are main effect terms for each list associated with the probability of being observed for the list, and θ_{ij}^{XY} is the interaction term which determines the magnitude and direction of dependency between the two lists. Estimating the model parameters allows one to estimate the number of victims not observed on any list, and thus the size of the hidden population. The approach allows for list interaction effects and ease in evaluating goodness-of-fit criteria based on summary statistics and visual plots. See Rivest and Daigle (2004) and Baillargeon and Rivest (2007) for the theoretical framework and empirical examples of such procedures when applied to commonly studied populations.

The common method to estimate parameters of such a Poisson log-linear model is through maximum likelihood. However, no overlap between administrative lists can be problematic and may result in what is known as "unidentifiability of the model" and "non-existence estimates for the model parameters". Chan et al. (2020) examine the non-existence of maximum likelihood estimators (MLEs) and unidentifiability constraints for such models. This is a commonly overlooked topic in the analysis of categorical data. In fact, most standard generalized linear modeling packages do not check for the existence of MLEs and when this problem exists, they report misleading estimates with large standard errors.

Chan et al. (2020) develop a model-fitting routine for sparse MSE data sets that is well-suited for population size estimation and which can handle existence issues. Essentially, the routine is a stepwise algorithm based on a predetermined threshold p-value. The algorithm commences with fitting a main effects model and then sequentially adds the most significant interaction terms one-by-one, provided that the resulting model passes non-existence of estimates and unidentifiable model checks. The algorithm is repeated until convergence to a final model. They apply their model-fitting routine to empirical data sets and find that it results in stable and reasonable estimates. An R package titled "SparseMSE" (Chan et al., 2019) has been developed and made publicly available for application of their methods to MSE data sets.

Addressing Parameter Redundancy and Model Unidentifiability

An issue related to non-existence and unidentifiability is parameter redundancy. Multiple list data can be displayed as a 2^m contingency table in which

m is the number of lists. Each variable has two levels (say 1 and 0) that respectively indicate whether individuals are or are not identified by a list. For example, for $m = 2$ with lists X and Y, the contingency table cell that corresponds to $X = 0, Y = 0$ contains the number of individuals that are not present in either list. As mentioned before, a standard model to fit to such count data is the Poisson log-linear model. However, this model may become parameter redundant and therefore unidentifiable because of the presence of possible zero cell counts in the table.

A parameter redundant model has parameters that are not estimable. We can follow a so-called parameter redundancy approach to obtain the subset of the original parameters that are estimable, as well as any estimable linear combinations of the original parameters. After detecting parameter redundancy, the original model can then be reparametrized as a smaller model with a smaller set of parameters that are all estimable. Those parameters have reliable estimates with reasonable standard errors. Examples of applying this method on ecological models can be found in Cole et al. (2010).

Catchpole and Morgan (1997) and Catchpole et al. (1998) describe a general method for detecting parameter redundancy for models that describe observations from distributions that belong to the exponential family of distributions, for example, Normal, Binomial or Poisson distributions. In this method, a derivative matrix is formed that contains the derivatives of means of table cell counts with respect to the log-linear model parameters. When the model is parameter redundant the rank of this matrix is smaller than the number of model parameters. The rank of the derivative matrix indicates the overall number of estimable model parameters and estimable functions of the parameters. All estimable parameters and linear combinations of them are obtained by solving a set of linear first-order partial differential equations (PDE).

Although a Poisson log-linear model is constructed to be identifiable, we expect this will not be the case after observing some zero cell counts. Sharifi Far et al. (2019) utilize the parameter redundancy method for Poisson log-linear models by adjusting the derivative matrix elements such that they include the observations, so any parameter redundancy caused by the number and position of observed zero cell counts is detected.

Assume fitting a Poisson log-linear model to a contingency table including some zero observations. If the rank of the derivative matrix equals the number of model parameters, then despite observing some zero cell counts the model is still identifiable. However, depending on the number and pattern of the zero cell counts, the model may become parameter redundant. In such a case some parameters are not estimable or only some linear combinations of them are estimable. Fitting the model under this scenario usually

shows large standard errors for estimates of parameters that are not directly estimable, indicating that these estimates are not reliable. There are other examples, in which the model is detected as parameter redundant but fitting it to the data with the specified pattern of zeros results in estimates with reasonable standard errors for all the parameters. Sharifi Far et al. (2019) explain that this happens because of existence of an "esoteric constraint" in the model. This constraint acts as an extra constraint on the parameters and together with the other estimable parameters of the model, makes all the model parameters estimable.

The approach described in section "Addressing Non-Existence of Maximum Likelihood Estimators and Model Unidentifiability," which is based on the work by Fienberg and Rinaldo (2012a, 2012b), detects identifiability of the model based on checking the existence of the MLEs. For a parameter redundant model, this method provides a subset of the initial parameters as the estimable parameters, but does not necessarily provide the estimable linear combinations of parameters. The parameter redundancy approach provides those estimable linear combinations, in addition to the esoteric constraint, when it exists. This process enables one to fit an identifiable log-linear model and obtain reliable estimates for the parameters to use in an MSE. Solving the relevant set of partial differential equations, as required, can be done in a symbolic algebra package, such as Maple (see Sharifi Far et al., 2019).

Matching When Linkages Are Not Directly Observed

Correctly linking data from different administrative lists is crucial for the successful implementation of multiple systems estimation. Some research areas such as clinical studies and epidemiology rely on unique subject identifiers. However, unique identifiers may not exist for some individuals identified by administrative lists relevant to hidden populations, as they are created by administrative bodies (for instance police, non-governmental organizations, or charities) for their own purposes.

The errors that occur when linking different lists are false-matches (linking records that belong to different individuals) and missed-matches (no linkage of records that belong to the same individual). Approaches to linkage are broadly either deterministic or probabilistic (Sayers et al., 2015). Deterministic linkage employs predetermined rules to effect matching. They are typically prone to missed matches, as errors (typographical or recording) can prevent matching records from the same individual. False matching is not observed frequently, as records are less likely to match exactly by chance. In

probabilistic linkage, a probability is assigned to every pair of records, with higher probabilities corresponding to more likely matches. Data are linked in accordance with some predetermined threshold. Probabilistic linkage is more prone to false-matches and less to missed matches. Alternative approaches include the use of Bayesian priors; see Goldstein et al. (2012). See also Harron et al. (2017) for more details on the above.

Missed-matches can result in bias, particularly when the error is non-random and depends on population subgroup. Bohensky et al. (2010) reported lower matching rates for subgroups according to age, sex, ethnicity and health status, which can translate to lower matches for vulnerable populations. Hence, considering the estimation of hidden populations, there is a need for observations to be recorded as precisely as possible.

False-matches can generate false associations or dilute true ones. See Harron et al. (2017) where several methods for evaluating the quality of linkage are described. A "gold standard" data set, where the true matches are known, may not be straightforward to obtain. Nevertheless, obtaining such a data set could assist in two ways. First, to evaluate the quality of the performed matching. Second, it could serve as a training set for informing Bayesian or machine learning algorithms that would perform probabilistic matching. In the absence of a "gold standard" data set, when multiple systems estimation is performed, data validation (identifying implausible scenarios within the data) and sensitivity analysis (by varying the threshold used in probabilistic matching) could be employed for quality evaluation. Bohensky et al. (2011) developed a series of reporting guidelines for studies involving data linkage. Recent work by Tibble et al. (2018) highlights the importance of including aliases in data linkage with vulnerable populations. The R package "RecordLinkage" (Borg and Sariyar, 2016) provides means to implement and evaluate different data linkage methods.

Covariate Information

Utilizing Covariate Information for Inference

Cruyff et al. (2017) summarize and apply an approach to population size estimation that is based on a Poisson log-linear model that incorporates categorical covariate information. Essentially, the observed counts of individuals corresponding to each possible capture history and covariate combination is regressed against the parameters corresponding to the lists upon which they are identified and observed levels of covariate information. This work is considered to be an extension of the Poisson log-linear models that have previously been introduced in the literature.

A forward stepwise approach is used to choose the most appropriate model upon which to base inference. Either the AIC or BIC criterion may be used. A parametric or nonparameteric bootstrap procedure is suggested by the authors in order to obtain standard errors and confidence intervals for the estimates. The authors discuss advantages and disadvantages of using these approaches.

This Poisson log-linear procedure has the benefit of directly estimating interaction effects between lists, between covariates, and/or between lists and covariates. Further, estimates of the sizes of subpopulations corresponding to specific covariate/demographic profiles can be obtained with ease, along with standard errors and confidence intervals. The utility of this procedure has been discussed in Cruyff et al. (2017) and has been instrumental in policy making decisions to combat human trafficking. An R package that can be used to recreate this work and to apply it to MSE data sets is in development.

A final note on the inclusion of covariates within the log-linear analysis concerns the Yule-Simpson paradox (Agresti, 2002). This is the phenomenon where the introduction of a third variable in the contingency table data may change the direction of the association between two existing categorical variables. In the context of a log-linear analysis it is possible that introducing a third variable in the analysis, with corresponding two- or three-way interactions, may change the sign of the estimated interaction effect between two pre-existing categorical variables.

Missing Covariate Information

It is not uncommon for administrative lists to have covariate information attached to each of the captured individuals. Such covariate information may come in the form of gender, age, and race. This information can generally be used to increase the efficiency of population size estimators and to obtain estimates corresponding to the subpopulations, as detailed in Cruyff et al. (2017).

In some cases covariate information may be missing or erroneously recorded for a subset of the captured individuals. When covariate information is used for the inferential procedure, the implications of missing data on the bias and variance of the estimators may be substantial. For such cases, a multiple imputation based approach to inference can lend itself well to account for the missing information.

Multiple imputation, as advocated by many researchers (Little & Rubin, 2002), is based on selecting an appropriate model for imputation. In the context of MSE, this would be based on the covariates and capture histories of

the observed individuals. The choice of imputation model is critical, and should be tested with techniques like cross validation.

At the inference stage, the missing information is repeatedly imputed to give a set of hypothetical full data sets, and Rubin's rules (Rubin, 1976) are used to obtain point estimates and standard errors. Conclusions can be drawn based on these estimates. The "mice" package in R (van Buuren, 2012) has the capability of performing multiple imputation on a wide range of data sets.

Local MSE Challenges

There are specific challenges that local MSE analyses may give rise to, relative to what is unlikely to be encountered with national MSE analyses. We discuss such anticipated challenges in this section.

For the local case, MSE data sets are likely to be based on administrative lists that come from regional law enforcement agencies or non-governmental agencies (NGOs) that operate in the area where the study population is situated. With respect to data collection, as such agencies typically operate independently (in contrast to the national case), it is unlikely there will be an agreed upon definition of the criteria required to identify individuals as part of the study population. This may give rise to lists which are either restrictive or relaxed toward the individuals they identify. That is, some lists may be restricted to only containing a subset of the study population (such as females), while others may contain individuals that fall outside the study population (such as sex workers that enter the sex trade business by their own accord). Hence, sparse overlap between the lists is likely to manifest. In such cases, it is imperative to collect as much covariate information as possible to assist in assessing the limitations of the study.

Challenges are likely to arise with obtaining permission to access such local lists. This may be due to ethical or confidentiality concerns. Further, anonymizing the information contained within lists may be resource intensive and is likely to add a burden to those organizations that are requested to provide the lists. A high-level of encouragement or incentives may be required by the study team for organizations to provide the lists.

With respect to inference, even if such anonymized lists are provided it could be the case that the quality of the data varies across the lists. For example, there may be erroneous entries and/or missing data that are unique functions of the lists, which can generate further difficulties in linking across lists and hence compound the difficulties that arise with sparse overlap. The methods discussed in the previous sections can assist with analyses for such cases.

Discussion

In this paper we have detailed several commonly encountered challenges when analyzing MSE data sets. These challenges, motivated by real data, arise from the data collection process in which there is a need for sharing of information across involved referral systems. Despite sharing information by the parties, non-overlapping lists are commonly observed. Adding covariate variables to usual MSE data is helpful due to providing extra information but can complicate the analysis. Moreover, some difficulties occur because of incorrectly linking different lists, or when the data come from local administrations rather than national ones. We have discussed methods and approaches that can be used to address these challenges.

Analyzing MSE data sets is especially challenging because the population consists of hidden individuals with erratic capture patterns. Further, a full set of direct observations on a human trafficking population to assess the performance of such methods may be nearly impossible. It is therefore important for rigorous MSE methods to be developed and made publicly available, while being upfront with the limitations of these methods. There is a growing number of R computing packages that can be used to analyze MSE data sets, as mentioned in this paper, when such challenges arise.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Kyle Shane Vincent https://orcid.org/0000-0002-3567-0798

Serveh Sharifi Far https://orcid.org/0000-0001-8403-6286

References

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). John Wiley & Sons, Inc.

Baillargeon, S., & Rivest, L.-P. (2007). Rcapture: Loglinear models for capture-recapture in R. *Journal of Statistical Software*, *19*, 1–31.

Bales, K., Hesketh, O., & Silverman, B. W. (2015). Modern slavery in the UK: How many victims? *Significance*, *12*, 16–21.

Bales, K., Murphy, L., & Silverman, B. W. (2018). How many trafficked people are there in New Orleans? Lessons in measurement. *Journal of Human Trafficking*, *6*(4), 375–387.

Bird, S. M., & King, R. (2018). Multiple systems estimation (or capture-recapture estimation) to inform public policy. *Annual Review of Statistics and Its Application*, *5*, 95–118.

Bohensky, M., Jolley, D., Sundararajan, V., Evans, S., Ibrahim, J., & Brand, C. (2011). Development and validation of reporting guidelines for studies involving data linkage. *Australian and New Zealand Journal of Public Health*, *35*(5), 486–489.

Bohensky, M., Jolley, D., Sundararajan, V., Evans, S., Pilcher, D. V., Scott, I., & Brand, C. A. (2010). Data linkage: A powerful research tool with potential problems. *BMC Health Services Research*, *10*(1), 346–352.

Borg, A., & Sariyar, M. (2016). *RecordLinkage: Record Linkage in R*. R package version 0.4-10.

Catchpole, E. A., & Morgan, B. J. T. (1997). Detecting parameter redundancy. *Biometrika*, *84*, 187–196.

Catchpole, E. A., Morgan, B. J. T., & Freeman, S. N. (1998). Estimation in parameter redundant models. *Biometrika*, *85*, 462–468.

Chan, L., Silverman, B. W., & Vincent, K. (2019). *SparseMSE: Multiple systems estimation for sparse capture data*. R package https://CRAN.R-project.org/package=SparseMSE

Chan, L., Silverman, B. W., & Vincent, K. (2020). Multiple systems estimation for sparse capture data: Inferential challenges when there are nonoverlapping lists. *Journal of the American Statistical Association*. Advance online publication. https://doi.org/10.1080/01621459.2019.1708748

Cole, D. J., Morgan, B. J. T., & Titterington, D. M. (2010). Detecting the parametric structure of models. *Mathematical Biosciences*, *228*, 16–30.

Cruyff, M., Overstall, A., & Papathomas, M. (in press). Multiple system estimation of victims of human trafficking: Model assessment and selection. *Crime & Delinquency*.

Cruyff, M., van Dijk, J., & van der Heijden, P. G. M. (2017). The challenge of counting victims of human trafficking: Not on the record: A multiple systems estimation of the numbers of human trafficking victims in the Netherlands in 2010–2015 by year, age, gender, and type of exploitation. *Chance*, *30*(3), 41–49.

Fienberg, S. E., & Rinaldo, A. (2012a). Maximum likelihood estimation in log-linear models. *The Annals of Statistics*, *40*(2), 996–1023.

Fienberg, S. E., & Rinaldo, A. (2012b). *Maximum likelihood estimation in log-linear models: Supplementary material* (Technical report), Carnegie Mellon University.

Goldstein, H., Harron, K., & Wade, A. (2012). The analysis of record-linked data using multiple imputation with data value priors. *Statistics in Medicine*, *31*, 3481–3493.

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M. L., & Goldstein, H. (2017). Challenges in administrative data linkage for research. *Big Data Society 4*(2), 2053951717745678.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*, (2nd ed.). Wiley Series in Probability and Statistics. Wiley.

Rivest, L.-P., & Daigle, G. (2004). Loglinear models for the robust design in mark-recapture experiments. *Biometrics*, *60*, 100–107.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*, 581–592.

Sayers, A., Ben-Shlomo, Y., Blom, A., & Steele, F. (2015). Probabilistic record linkage. *International Journal of Epidemiology*, *45*, 954–964.

Sharifi Far, S., King, R., Bird, S., Overstall, A., Worthington, H., & Jewell, N. (2020). Multiple systems estimation for modern slavery: Robustness of list omission and combination. *Crime and Delinquency*. Advance online publication. https://doi.org/10.1177/0011128720951429

Sharifi Far, S., Papathomas, M., & King, R. (2019). Parameter redundancy and the existence of maximum likelihood estimates in log-linear models. *Statistica Sinica*. Advance online publication. https://doi.org/10.5705/ss.202018.0100

Silverman, B. W. (2020). Multiple systems analysis for the quantification of modern slavery: Classical and bayesian approaches. *Journal of the Royal Statistical Society, Series A*. *183*(4), 691–736.

Tibble, H., Di Law, H., & Spittal, M. E. A. (2018). The importance of including aliases in data linkage with vulnerable populations. *BMC Medical Research Methodology*, *18*, 1–5.

van Buuren, S. (2012). *Flexible imputation of missing data*. Chapman and Hall/CRC Press.

Williams, B. K., Nichols, J. D., & Conroy, M. (2002). *The analysis and management of animal populations*. Academic Press.

Author Biographies

Kyle Shane Vincent is an independent consultant whose research focuses on developing innovative strategies that can be used to efficiently study hard-to-reach populations. Much of his work focuses on using network sampling strategies and mark-recapture procedures, which is motivated by challenges encountered when studying populations such as those comprised of human trafficking victims, drug users, or commercial sex workers. Dr. Vincent received his PhD from Simon Fraser University.

Serveh Sharifi Far is a university teacher in Statistics in the School of Mathematics at the University of Edinburgh. Her research interests include parameter redundancy and analysis of categorical data motivated by social science and biological data. She received her PhD from the University of St Andrews.

Michail Papathomas is a lecturer in Statistics at the School of Mathematics and Statistics at the University of St Andrews. His research is on model comparison/variable selection for exploring the presence of interactions between variables and their complex association with some outcome. His research also concerns identifiability for log-linear models fitted to contingency tables. Dr Michail Papathomas received his PhD from the University of Sheffield, UK.